ATTACHMENT B

Description of SM Clustering Methodology
Using the Denver Curtis Park Wire Center (DNVRCOCP)

Step 0:

The customer geo-coded data is read from the *.in file. The point coordinates are converted from longitude/latitude to relative distances from the central office. Dividing the rectangle that encloses the wire center into squares that are 500 ft. on a side form a starting set of microgrids. The points are then assigned to a specific microgrid. The centroid of the microgrid is calculated using line weighting. If the number of populated microgrids is less than 3000 then the process is complete and the clustering begins. If the number of populated microgrids is greater than 1500 then the length of the microgrid is increased by 500 ft and the process is repeated.

Prior to forming the microgrids, the customer line counts are adjusted to match the entries in the line count table of the database. This includes adjusting the business lines to include public telephone lines and special access channels.

During this process the maximum distance within a microgrid is also calculated. This distance is used to reduce the client input maximum copper distance.

For DNVRCOCP, the statistics were approximately:

      14,260 microgrids
        760 populated microgrids
        500 ft microgrid size
        701 ft maximum distance within a microgrid

Step 1:

The microgrids are roughly divided into two groups. Those that are within a copper service area and those that should be served with DLC. The copper customers are in cluster 1 and the DLC customers are within cluster 2. For DNVRCOCP, the effective copper distance limit from the central office for inclusion in cluster 1 is 16,299 feet.

At this stage there are no constraints on line counts or distances, other than the maximum copper distance is reduced by the maximum distance within a microgrid.

Step 2:

At this stage the microgrids in cluster 2 are divided. The new clusters are constrained to meet the distance constraints and also the line limit constraints of a Serving Area Interface (SAI). The line limits are calculated from the adjusted customer line data discussed in Step 0. Default values are 1800 lines with an 80% fill, which yields an effective line constraint of 1440.

These step produces non-contiguous clusters (see clusters 3 through 7 and cluster 9).

Step 3:

This is the 'simple reassignment' phase of the clustering process. In this step, the microgrids are reassigned to the nearest cluster. The cluster centroids are retained from Step 2. There are additional constraints placed on the clustering at this point. For some reason, if a microgrid is to move it must be 1.5 times closer to the new cluster centroid than to the current centroid.

As can be seen from the charts, this step removes the disjointed pockets of customers and non-contiguous clusters left in the Step 2.

Step 4:

This if the 'full optimization' step. If the number of microgrids is less than 1000, then this supposed optimization takes place. It is an attempt to reduce or "tighten up" the clusters. In reality it has little or no effect given the steps that follow.

Step 5:

In this step the clusters with large numbers of lines are split into multiple clusters. The attempt is made to reduce the number of lines to meet the line limit constraint. At this step the client input fill factor is changed to 1.0, effectively increasing the line limit to 1800 lines.

Step 6:

Simple reassignment is applied yet again. See Step 3 above.

Step 7:

The results are written out to the *.clu file. The first portion of this file is a summary of line counts, coordinates and other information by cluster. Before this summary information is written out, each cluster is restored to the original line counts.

Below is a summary of the line counts for each of the clusters for each processing step:

Cluster	Processing Step						
	1	2	3	4	5	6	7
1	84545	84545	66282	66282	16069	16093	7925
2	10756	718	3548	3548	1758	1774	897
3		1439	4527	4527	929	1105	798
4		1438	2845	2845	1045	1026	871
5		1439	4153	4153	600	1308	855
6		1437	3830	3830	288	1552	791

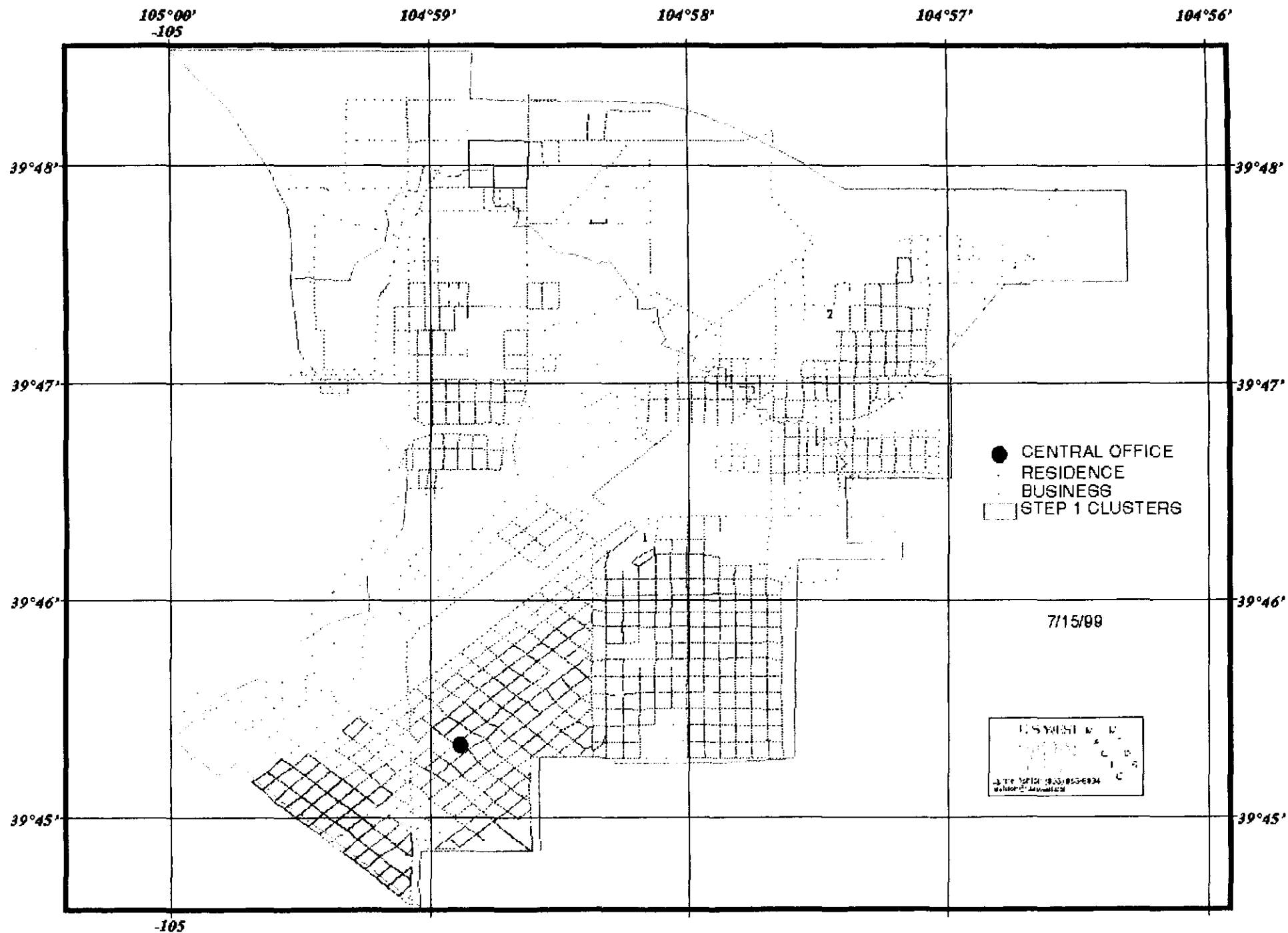|							
7		1439	2326	2326	529	1379	700
8		1438	3354	3354	1559	1528	772
9		1406	4444	4444	890	1700	955
10					1799	1746	889
11					1796	1617	1169
12					1789	1338	846
13					1799	525	397
14					1735	863	456
15					1779	754	432
16					1635	3287	1766
17					1644	2383	1272
18					1704	685	645
19					1319	1676	1042
20					1534	2550	1336
21					1440	1440	732
22					1032	1032	512
23					1490	1889	944
24					7444	7926	3934
25					2052	2052	1065
26					1002	1091	545
27					1687	1687	833
28					1856	2133	1069
29					3963	3963	1957
30					9704	9704	4780
31					1789	1399	984
32					1799	1674	1478
33					1798	1178	1051
34					1799	1569	886
35					1798	1717	849
36					1764	1770	1011
37					1796	1049	673
38					1735	1607	924
39					1797	852	421
40					1795	1752	899
41					1797	1791	881
42					1756	1119	596

Problems:

1. The cluster process works with the microgrid centroid locations rather than with the customer location directly. The formation of microgrids is not necessary because the clustering process can work directly with customer location data. Apparently, the cluster process applied in this section of the SM is a holdover from some earlier model that was constrained by memory or processing environment.
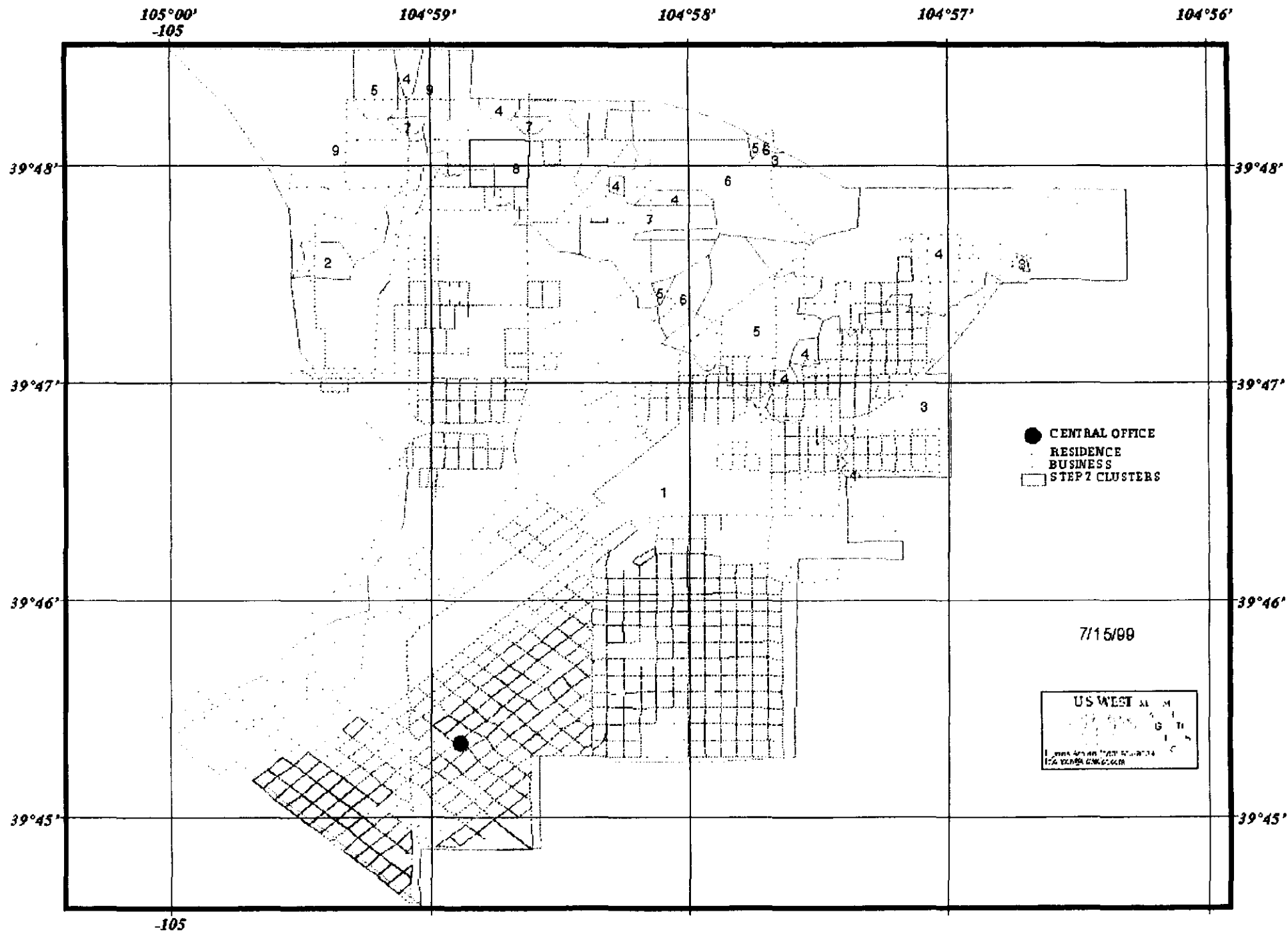
2. The adjustment to the maximum copper distance is arbitrary. The maximum distance in a given microgrid, point to microgrid centroid, has little or no bearing on the cable distance. The distance constraint for the clustering process needs to be calculated individually for each point in the microgrid.
3. There are problems with the 'simple reassignment' code. If a cluster has to many points and its nearest neighbors are also beyond limits, then nothing is done. At the very least a new cluster should be formed.
4. All of the line adjustments that take place within the model should be consolidated. It is very confusing, apparently even to the SM modelers, how the lines are adjusted. As a matter of fact, the CLUSINTF process should be entirely eliminated. It would make much more sense to consolidate the CLUSTER and CLUSINTF processes in one process.
5. There is no recognition in the SM clustering of the wire center boundaries. Looking at cluster 4 on chart 6 you can see that the distribution plant for this cluster will leave the wire center boundaries. It is also apparent that there is a strong possibility that the feeder plant would escape the wire center boundaries.
6. The division of customers into copper/DLC areas is wasted. When the 'simple reassignment' step, Step 3, is completed any distinction between copper and DLC is gone. To be done properly, the customers served on copper and the customers served with DLC need to be treated separately. This would call for a redesign of the SM clustering process.
7. Changing the fill factor in mid-process is flawed. Whatever factor is used, it should be used consistently. By changing the factor in mid-process, the number of clusters generated is reduced thus reducing the total loop cost.
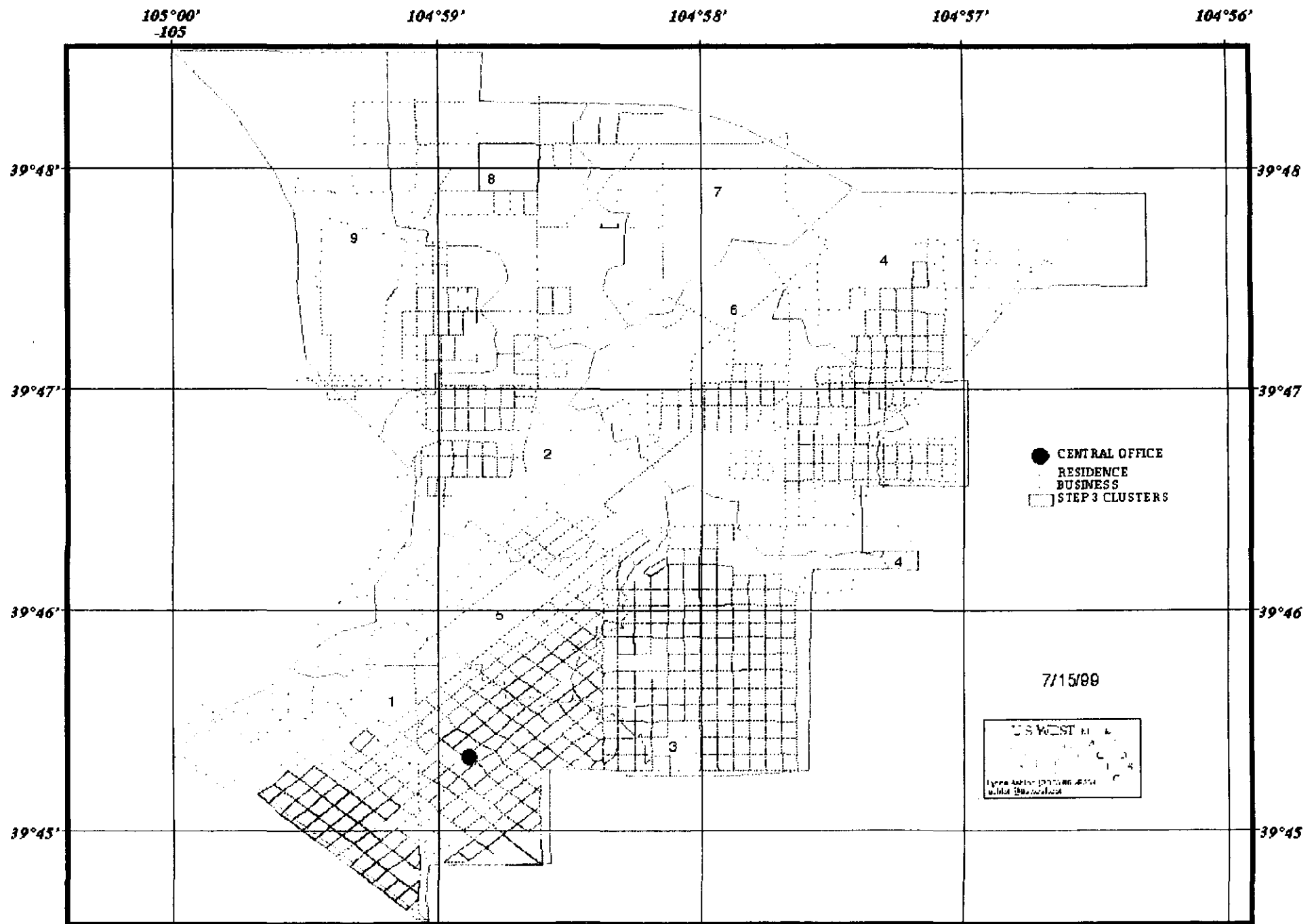
ATTACHMENT C

DENVER CURTIS PARK DATA - STEP 1 - FIRST TWO CLUSTERS

● CENTRAL OFFICE
  RESIDENCE
  BUSINESS
  ☐ STEP 1 CLUSTERS

7/15/99

DENVER CURTIS PARK DATA - STEP 2 - INITIAL CLUSTERS

105°00'
-105

104°59'

104°58'

104°57'

104°56'

39°48'

39°47'

39°46'

39°45'

● CENTRAL OFFICE
· RESIDENCE
BUSINESS
☐ STEP 2 CLUSTERS

7/15/99

US WEST

-105

DENVER CURTIS PARK DATA - STEP 3 - SIMPLE REASSIGNMENT

● CENTRAL OFFICE
  RESIDENCE
  BUSINESS
☐ STEP 3 CLUSTERS

7/15/99

DENVER CURTIS PARK DATA - STEP 4 - FULL OPTIMIZATION

CENTRAL OFFICE
RESIDENCE
BUSINESS
STEP 4 CLUSTERS

7/15/99
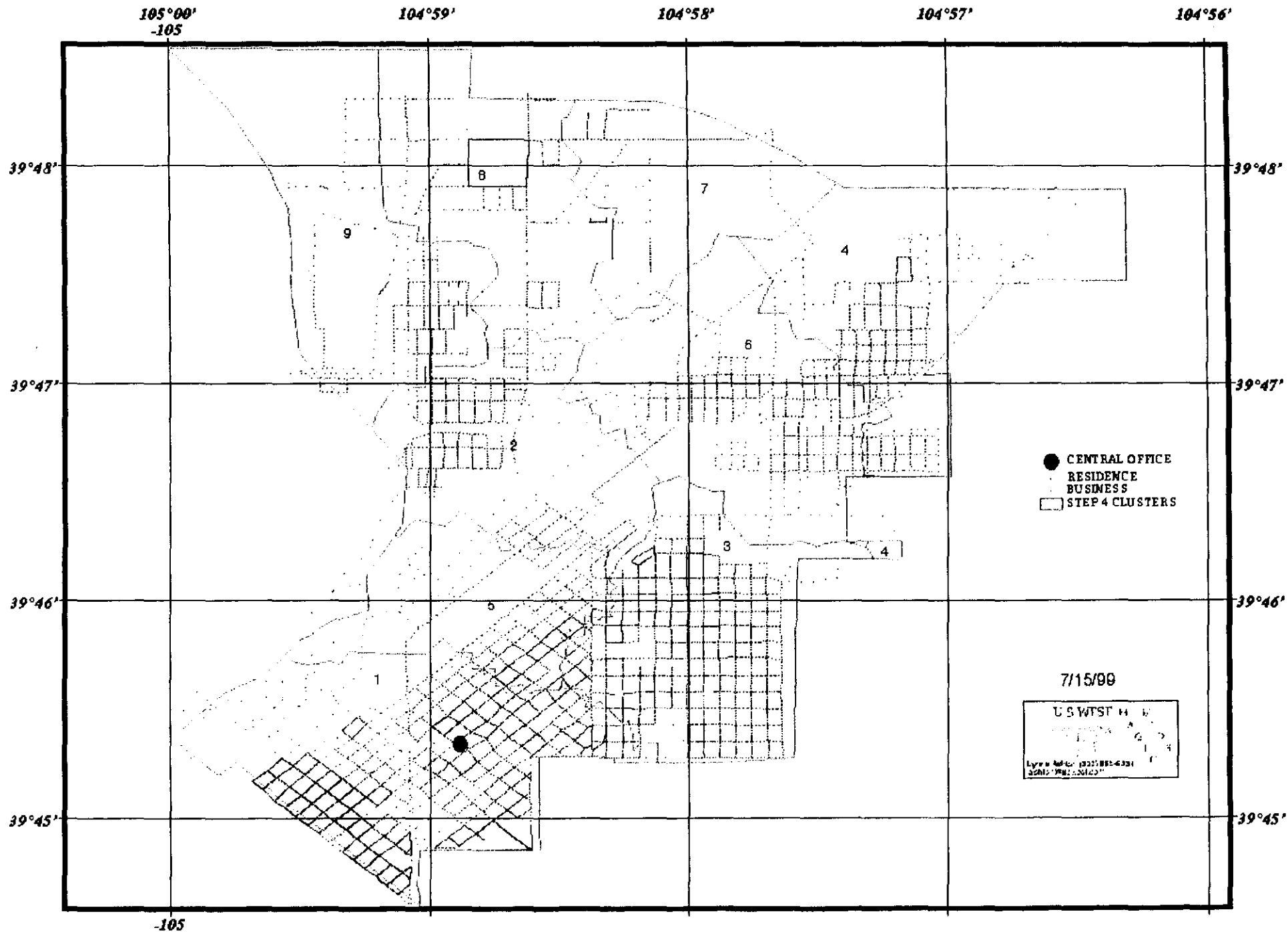
DENVER CURTIS PARK DATA - STEP 5 - BREAK UP LARGE CLUSTERS

CENTRAL OFFICE
RESIDENCE
BUSINESS
STEP 5 CLUSTERS

7/15/99
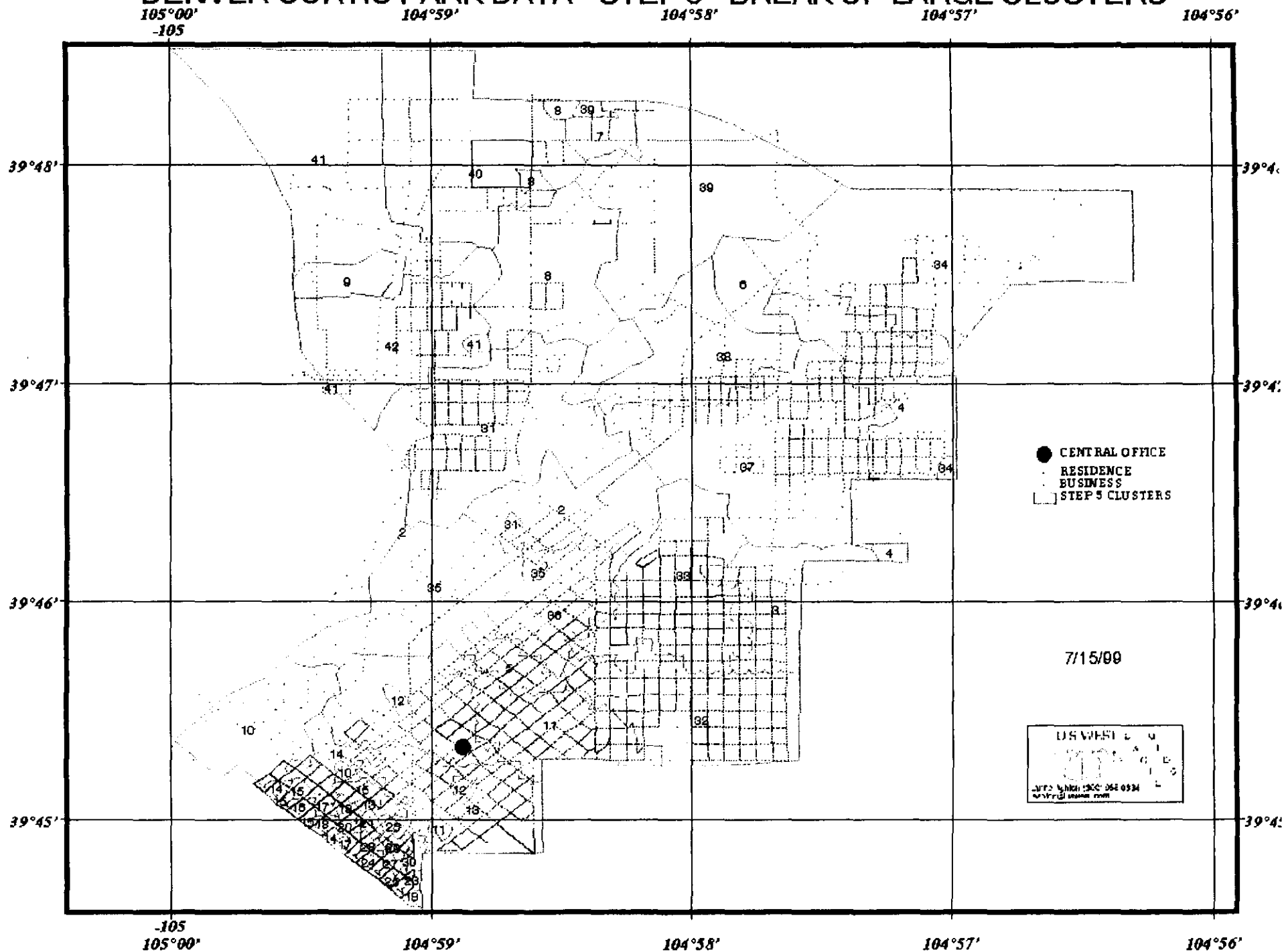
DENVER CURTIS PARK DATA - STEP 6 - LAST SIMPLE REASSIGNMENT

105°00'
-105
104°59'
104°58'
104°57'
104°56'

39°48'

40
7

39°48'

41

39
39

34

8

5

6

39°47'

42

39°47'

38

31

4

37

2

● CENTRAL OFFICE
· RESIDENCE
  BUSINESS
☐ STEP 6 CLUSTERS

4

35
36

4

3

39°46'

39°46'

5

33

7/15/99

14

32

US WEST

10

12

11

15

13

16

17
19

16

18

20
21
26

22
25

24
37
23

1

105°00'
104°59'
104°58'
104°57'
104°56'